

# THE CMCC-INGV GLOBAL OCEAN DATA ASSIMILATION SYSTEM (CIGODAS)

By **Pierluigi Di Pietro**

Istituto Nazionale  
di Geofisica e Vulcanologia  
(INGV)  
[dipietro@bo.ingv.it](mailto:dipietro@bo.ingv.it)

and **Simona Masina**

Centro Euro-Mediterraneo  
per i Cambiamenti Climatici  
(CMCC)  
Istituto Nazionale  
di Geofisica e Vulcanologia  
(INGV)  
[masina@bo.ingv.it](mailto:masina@bo.ingv.it)

**SUMMARY** This report summarizes the technical structure and main characteristics of the CMCC-INGV Global Ocean Data Assimilation System (CIGODAS) based on a Reduced Order Optimal Interpolation scheme and a coarse resolution Global Ocean Model for the assimilation of temperature and salinity observations. It is intended to be a reference guide for new users who are interested in setting up and running an experiment using this approach and producing estimates of the time-varying, three-dimensional state of the global ocean.

**Keywords:** Oceanic General Circulation Models, Ocean Analysis, Data Assimilation, Optimal Interpolation

---



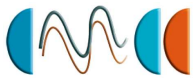
THE CMCC-INGV GLOBAL OCEAN DATA ASSIMILATION SYSTEM (CIGODAS)

# 02

Centro Euro-Mediterraneo per i Cambiamenti Climatici

# Contents

Introduction . . . . .	4
Global Ocean Data Assimilation System: components description . . . . .	4
Ocean Model . . . . .	4
Assimilation scheme . . . . .	5
Observational data . . . . .	9
Compiling environment . . . . .	10
Running environment . . . . .	13
Auxiliary files . . . . .	15
Single Profile Assimilation Experiment: sensitivity to EOF . . . . .	21
<b>Appendix 1: OGCM namelist (OPA 8.2)</b>	<b>25</b>
<b>Appendix 2: Directory tree</b>	<b>37</b>



## INTRODUCTION

The CIGODAS system is composed of the Ocean General Circulation Model OPA 8.2 [Madec et al., 1999] in the ORCA2 global configuration ( $2^\circ$  longitude x  $2^\circ \cos \varphi$  latitude), and an Optimal Interpolation (OI) scheme based on the System for Ocean Forecasting and Analysis (SOFA) assimilation software [De Mey and Benkiran, 2002]. The original SOFA code has been implemented to the global ocean for the assimilation of temperature and salinity [Bellucci et al., 2007]. This report describes this global ocean assimilation system and some recent developments aimed to produce multi-decadal analyses of the 3-D ocean state, and initial conditions for seasonal forecasting systems.

## GLOBAL OCEAN DATA ASSIMILATION SYSTEM: COMPONENTS DESCRIPTION

### OCEAN MODEL

The OGCM is OPA 8.2 [Madec et al., 1999] in ORCA2 configuration with the following characteristics.

- Grid and numerical scheme :

- horizontal resolution:  $2^\circ \times 2^\circ \cos \varphi$  Mercator grid everywhere, except for the  $20^\circ\text{N}$ - $20^\circ\text{S}$  latitude belt, where the meridional grid spacing is progressively decreased to  $0.5^\circ$  to better resolve the equatorial dynamics. Quasi-isotrope tri-polar grid (2 poles in the northern hemisphere, one over Canada and the other over Siberia).
- Vertical resolution: 31 vertical levels with 20 levels lying in the top 500 meters.

- Advection scheme : 2nd order Arakawa
- time-stepping: leap-frog except for lateral diffusion (forward) and vertical diffusion (backward)
- Vertical coordinate : z-coordinate with free surface.

- Major parameterizations:

- The vertical mixing parameterization is based on a turbulent kinetic energy prognostic equation [Blanke and Delecluse, 1993].
- Tracers are diffused along Isopycnal surfaces with an eddy mixing coefficient of  $2000 \text{ m}^2 \text{ s}^{-1}$ , using a 2nd order operator.
- The horizontal viscosity coefficient is space dependent: a typical value of  $40\,000 \text{ m}^2 \text{ s}^{-1}$  is applied poleward of the  $20^\circ\text{N}$ - $20^\circ\text{S}$  belt, while it gradually decreases to  $2000 \text{ m}^2 \text{ s}^{-1}$  in the equatorial region, except for the western boundary regions (downgradient scheme). These coefficients are used in the tracer diffusion equations (2nd order operator).
- Lateral diffusion is supplemented with the Gent and McWilliams (1990) eddy-induced velocity parameterization [Lazar et al. 1999]. A typical strength of the eddy-induced velocity is  $2000 \text{ m}^2 \text{ s}^{-1}$ , but it is decreased in the equatorial region.

- The time step is 1 h 36 min for both dynamics and tracers.
- There is no sea ice model coupled to the ocean model. The ice cover is evaluated comparing the local sea surface temperature to a surface freezing potential temperature defined by a Unesco formula [

UNESCO, 1983].

- The model bottom topography is defined as steps that follow the face of the model cells (steplike topography) [Madec et al. 1996]. The model bathymetry is constructed from the ETOPO 5'x5' global bathymetric field.
- Air-sea boundary conditions:
  - Momentum flux are prescribed at surface as the boundary condition on the vertical turbulent momentum fluxes (compiling option Key\_tau )
  - Heat and Fresh Water Fluxes are prescribed at surface as boundary conditions on the vertical turbulent fluxes on T and S (compiling option key\_fluxforced )
  - Penetrative sunlight penetration is implemented using 2 extinction lengths [Blanke and Delecluse 1993] (compiling option Key\_fluxqr)
- The model doesn't implement tidal mixing and river mouth mixing.
- Lateral boundary conditions are set to zero fluxes of heat and salt, while velocities tangential to the coast are set to zero (no-slip).

A complete description of all possible physical parameterization of the model can be found in the OPA reference manual [Madec et al 1999] available on the web (<http://www.lodyc.jussieu.fr/opa/>).

## ASSIMILATION SCHEME

The assimilation scheme is a Reduced Order Optimal Interpolation (ROOI) scheme implemented to produce local corrections to the

temperature and salinity fields. The data assimilation system used to produce the present ocean reanalyses is based on SOFA [De Mey and Benkiran, 2002] and has been designed for the assimilation of in-situ vertical profiles of temperature and salinity. OI is a particular case of the extended Kalman filter, where the forecast error covariance matrix is replaced with a background error covariance matrix  $B_f$  [Daley, 1991]:

$$B_f = (D_f)^{1/2} C (D_f)^{1/2} \quad (1)$$

In eq (1) correlations  $C$  are assumed to be constant in time while the background error variance diagonal matrix  $D_f$  is generally predicted using an external scheme from the analysis error variances at the previous filter step [Ide et al. 1997] or modeled through empirical multivariate structure functions [De Mey and Benkiran, 2002]. The analysis  $x^a$  at each time step is computed as

$$x^a = x^f + K^{OI} [y^o - H(x^f)] \quad (2)$$

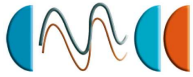
with  $x^f$  indicating the state forecast,  $y^o$  the observations,  $H$  the observation operator, and  $K^{OI}$  the OI Kalman gain defined as

$$K^{OI} = B^f H^T (H B^f H^T + R)^{-1} \quad (3)$$

where  $R$  is the observation error covariance matrix.

The order reduction is achieved by projecting the state vector onto a set of precalculated vertical empirical orthogonal functions, composing the columns of an  $S$  matrix (the simplification operator). The vertical EOFs are the eigenvectors of the  $B^f$  matrix that is factorized as follows:

$$B_f = S^T B_r^f S \quad (4)$$



with  $B_r^f$  containing the horizontal correlations of the multivariate EOF modes defined in  $S$ . After substituting eq (4) into eq (3) we obtain

$$K^{OI} = S^T B_r^f H_r^T (H_r B_r^f H_r^T + R_r)^{-1} \quad (5)$$

where  $H_r = H S^T$  and the matrix  $R_r$  takes into account the representativity error in the reduced space.

By defining the OI Kalman gain in the reduced space as

$$K^{ROOI} = B_r^f H_r^T (H_r B_r^f H_r^T + R_r)^{-1} \quad (6)$$

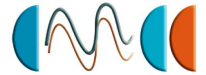
eq (5) can be simply written as

$$K^{OI} = S^T K^{ROOI} \quad (7)$$

## ROOI MAIN PARAMETERIZATIONS

The ROOI scheme is implemented through the System for Ocean Forecasting and Analysis (SOFA) assimilation software [De Mey and Benkiran, 2002]. This software is tunable by modifying the values of namelist parameters (table 8 and 9), which are read at startup. The most important parameters are related to the domain geometry, the spatial and temporal shape of corrections, and the number of observations to be taken into account in every single correction process and which affects the computational time :

- **NMASK**: The number of regional masks to be defined for diagnostic calculations. Standard value for experiments where gridpoint defined eofs (V2 or V3) are used is NMASK=27118 (being 27148=149\*182 the horizontal resolution in ORCA2 configuration).
- **RGY, RGY**: The guess error correlation radii for zero time lag in the x and y direction in kilometers. In our experiments, RGX=RGY=300 Km. This affects the spatial shape of the gaussian weighing function applied to the innovations.
- **RGT**: The guess error correlation radii in time. In our experiments, RGT=7 days. This affects the temporal shape of the gaussian weighing applied to the innovations.
- **ROX, ROY**: The observational error correlation radii for zero time lag in the x and y direction in kilometers. In our experiments, we consider errors on observations spatially uncorrelated, setting ROX=ROY= $1e^{-6}$  Km.
- **ROT**: The observational error correlation radii in time (in days). In our experiments, we consider errors on observations temporally uncorrelated, setting ROT= $1e^{-6}$  days
- **RIX, RIY, RIT**: The selection radii for zero time lag in the x and the y directions (in kilometers) and in time (in days). The selection radii define a perimeter or "bubble" of generally elliptic shape in space-time, within which the data-model misfits are inverted around each point of the analysis grid. In our experiments, RIX=RIY=800 Km. and RIT=7 days.
- **NSEL**: The maximum number of influential points which are to be selected from each dataset for suboptimal objective analysis. The algorithm selects at most NSEL data in the bubble which are 'less' correlated with each other than with the interpolation point (this behavior can



be refined by REDUN). In our experiments NSEL=20 has been chosen to balance correction accuracy with computational time.

- REDUN: When selecting data in the influence bubble, the suboptimal algorithm eliminates from the selection observations which are closer than  $1/\sqrt{REDUN}$  times their distance to the interpolation point. Increasing this value will have the effect of keeping more observations for a given influence radius, and thus of increasing the observational density and informational redundancy around the interpolation point. In our experiments, we take the default value REDUN=1.0.
- NRED: The local size of the reduced space, i.e. the number of modes to be estimated by SOFA. In our experiments, the size of reduced space is set to NRED=10.
- IFCONV: The type of conversion between the data coordinate system and the analysis/model canvas coordinate system. Possible values of IFCONV are
  - 0 :everything is on a cartesian frame. This choice applies to process models
  - 1 data are located in geocentric coordinates (lon,lat). The canvas is cartesian. The data are projected onto the plane that is tangent at the origin of the canvas. This is sometimes called 'beta-plane projection'. This choice applies to open-boundary mesoscale models
  - 2 data are located in geocentric coordinates (lon,lat). The canvas is also defined in geocentric coordinates. At each gridpoint, the data are combined on the local tangent

plane. This choice applies to basin-scale ocean models

- 3 same as 2 but the canvas is periodic in X. In this case, the last row of canvas points is assumed to be connected to the first row. This choice applies to global ocean models

Our choice, given the structure of the ORCA2 coordinates (from 180W to 180E, periodic in x, but the periodicity meridian is located at 70E, so the longitudes can be considered 'open' at 180W - 180E ), is IFCONV=2.

- TTMIN: The starting date of the analysis. All times in SOFA are assumed to be Julian Dates. This parameter is set by the job script, which evaluates the date using a calendar algorithm.
- NTIMES: The number of integration/analysis cycles. TTMIN=2 in our experiments: first cycle is dedicated to the generation of innovations and corrections, the second is limited to the storing of the correction in the restart file.
- TINT: The interval between estimates. In our experiments, TINT=7 days.
- IFHCST: Chooses how misfits in the past and future (hindcast) will be handled. This parameter has two values:
  - 0: filter mode. Only data misfits prior to the analysis time are used in the correction. Misfits are differences between the data and the model at the same time as the observations.
  - 1: smoother mode. Data misfits within RIT time distance of the analysis time (both in past and in future) are used. This requires running the

model some time in the 'future' in order to calculate the model-data misfits.

An important feature of this ROOI scheme lies in the multivariate structure of the background error covariance matrix, which spreads the corrections over different types of variables. In the present implementation the state vector is defined as the bivariate temperature and salinity vector. This implies that when, for example, only vertical temperature profiles are assimilated, corrections are applied to salinity as well by the vertical EOFs.

The multivariate EOFs used for assimilating in situ data have been diagnosed from the synthetic dataset of vertical temperature and salinity profiles provided by an ocean model simulation where no data assimilation has been applied [Belluci et al. , 2007]. Only the first ten dominant modes are retained (order reduction) [Sparnocchia et al, 2003].

In version V1 of multivariate EOFs , they are calculated for different sub regions (fig. 1), each representing different dynamical regimes, and are time independent.

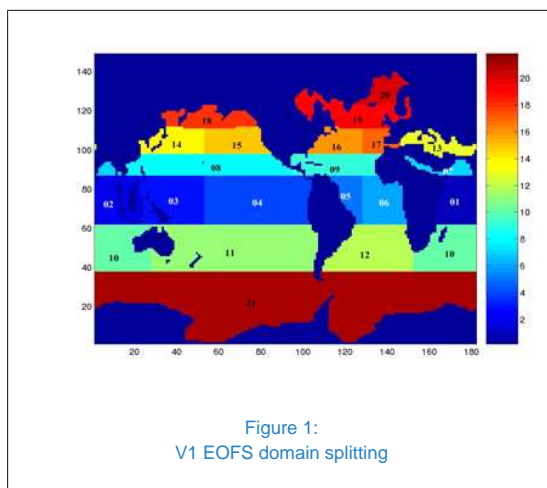


Figure 1:  
V1 EOFs domain splitting

In version V2 of multivariate EOFs, they are calculated for each grid point, and are seasonally

dependent (fig. 2). The Eofs are evaluated from horizontally smoothed Temperature and Salinity fields, using a three point radius moving average.

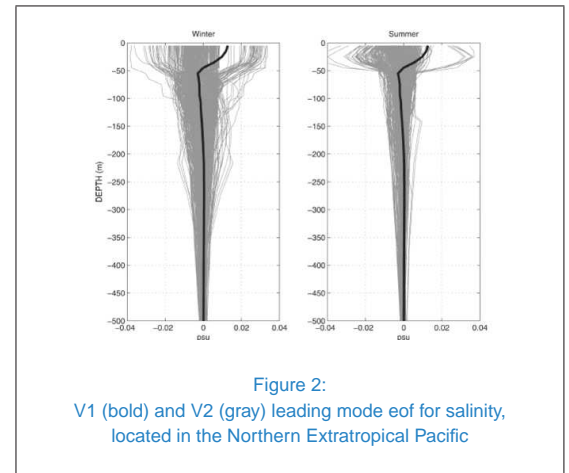


Figure 2:  
V1 (bold) and V2 (gray) leading mode eof for salinity,  
located in the Northern Extratropical Pacific

In version V3 of multivariate EOFs, they are calculated for each grid point, and are seasonally dependent. The Eofs are evaluated from Temperature and Salinity fields without horizontal smoothing.

## EOFS BUILDING PROCEDURE

The eofs are evaluated using the standard MatLab(tm) pcacov function. This function perform Principal Components Analysis (PCA) using as input a Covariance Matrix.

We use as source the monthly Temperature and Salinity fields produced by our OGCM in which no assimilation is performed. The monthly time-serie starts from January 1962 and ends at December 2001 (480 months).

To evaluate the seasonal EOFs in each gridpoint, a subset (January-March for Winter, April-June for Spring, July-September for Summer, October-December for Fall) of monthly-mean T,S profiles is extracted (120 months)from the main timeserie . Every single temperature and salinity profiles are merged



in a 62 element vector, called the bi-variate vertical sequence.

A seasonal bi-variate timeserie is a matrix  $A$  with size  $62 \times 120$ , in which rows 1 to 31 are temperature and rows 32 to 62 are salinity. Along the columns is the time dimension.

The seasonal mean  $M$  and Standard Deviation  $SD$  (two  $1 \times 62$  arrays) are then evaluated from this matrix. The standard deviation array collects the temperature and salinity standard deviation evaluated at each level. An average standard deviation, evaluated as a vertical mean of the 31 different standard deviations, is computed for both temperature  $SD_t$  and salinity  $SD_s$ .

The seasonal mean  $M$  is then removed from each column of  $A$ , obtaining a  $62 \times 120$  matrix of anomalies  $H = A - M$ . The Temperature-Salinity anomalies are then normalized separately dividing them by the respective averaged standard deviations  $SD_t$  and  $SD_s$ ,  $H_{n_{i=1,31}} = H_{i=1,31}/SD_t$ ,  $H_{n_{i=32,62}} = H_{i=32,62}/SD_s$  obtaining the  $n$ -dimensional normalized matrix  $H_n$ .

The `pcacov` function is then applied to the covariance matrix  $C = (H_n * H_n^t)/(m - 1)$ , being  $m$  the time dimension length of the  $A$  matrix (120 months). obtaining as output the eigenvectors  $V$ , the eigenvalues  $eigv$  and the related explained variance  $exv$ . The eigenvectors are then re-dimensionalized multiplying them by the standard deviation, i.e.,  $V_{i=1:31} * SD_t$ ,  $V_{i=32:62} * SD_s$ . The first ten modes are then stored into a file in this order:

- the regional mask number (for gridpoint eofs, we have  $149 \times 182$  'regional masks')
- the square root of the  $n^{th}$   $eigv$

- the  $n^{th}$  eigenvector

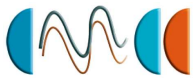
This procedure is repeated for the V2 and V3 EOFs for every grid point (considered as regional masks) on the horizontal domain.

The procedure is similar for the V1 EOFs set: the only differences being that, in every subdomain, the bi-variate timeserie matrix  $A$  is built using T and S vertical profiles previously horizontally averaged over the subdomain itself, and all times are taken into account (480 months).

## OBSERVATIONAL DATA

The CIGODAS has been tested only for the assimilation of vertical profiles of temperature and salinity, on the global domain. The assimilated profiles come from the ENSEMBLE quality controlled subsurface ocean temperature and salinity dataset [<http://hadobs.metoffice.com/en3/>]. Such dataset is built merging different source data, as WOD05 [[www.nodc.noaa.gov/OC5/WOD05/pr\\_wod05.html](http://www.nodc.noaa.gov/OC5/WOD05/pr_wod05.html)], GTSP [[www.nodc.noaa.gov/GTSP/gtspp-home.html](http://www.nodc.noaa.gov/GTSP/gtspp-home.html)] and Argo [[www.argo.net/](http://www.argo.net/)]. The original data are quality checked by the MetOffice staff [Ingleby and Huddleston, 2007] and finally delivered. The data are available in NetCDF format. To be used by the model, data are further processed:

- only potential temperature and salinity data with quality flag equal to 1 (all checks passed) is retained
- the vertical profiles are interpolated onto the t-grid levels, further decimating the data
- the resulting interpolated profiles are stored in monthly files, in ascii format.



- each file has a header, composed of a 62-element array, containing the estimated background error (set to all levels to 0.5 C for temperature, and 0.5 or 0.2 PSU for salinity, depending on the experiment)
- each profile entry is composed by:
  - a five element array defining the observation type (0 for indirect scalar data), longitude, latitude, time index and the number of observations
  - an array containing the position of each profile observation in the bivariate state vector (from 1 to 31 for temperature, from 32 to 62 for salinity data)
  - the profile observed values (temperature or salinity)
- Each monthly file (except December) has profiles of the entire next month appended. The profile time axis origin is set to the beginning of current month (for example, the March file contains the March profiles indexed from 1 to 31, and the following April profiles indexed from 32 to 61; the April file contains April profile indexed from 1 to 30 and May profiles from 31 to 62 and so on). The SOFA code sets the time origin at the beginning of the starting month, then select the data using this time as reference. When the time window of an assimilation cycle is located over the end of a month and the beginning of the next, it is necessary to have the profiles of subsequent month in the right time axis, to perform correct assimilation.

## COMPILING ENVIRONMENT

The model is designed to run on NEC-SX9 vector super computer and the compilation can be done natively on the NEC-SX9 or using

NEC cross-compiler on machines mounting the NEC-SX9 file system.

To compile the code:

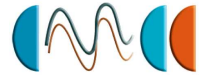
1. log in to the system
2. go to the `.../CIGODAS/work` directory where the `cpl_sofa_qu` shell file is located. There are four compiling options. The preferred one has to be edited assigning the right value to the "RUNNAME" string variable. "VERT\_PROF\_ASSIM" string is used as the default value for "RUNNAME". It is designed for vertical profile assimilation, and it is the only one that has been implemented in the global ocean. Other available values are: "SLA\_ASSIM" for sea level assimilation; "INTER" for interannual forcing run; "PERP" for perpetual year run, but have never been tested at INGV-CMCC.
3. Invoking `cpl_sofa_qu` using the `sh` shell interpreter, some environment variable are automatically set, like the "BASEDIR" variable that defines the root directory of the entire model. A number of Makefiles are then sequentially invoked using the `sxgmake` interpreter.

In every *Makefile* the following NetCDF variables are to be set, in order to link the correct version of NetCDF libraries and include files:

- `NCDF_INC = /SX/lib/netcdf-3.6.1-ew/include`
- `NCDF_LIB = /SX/lib/netcdf-3.6.1-ew/lib -lnetcd`

The Makefiles called in sequence are:

- the IOIPSL Makefile, which compiles the source files of the input/output modules, and produce a set of I/O



libraries used by the ocean component

- the SOFA Makefile, which compiles the SOFA source code and produce a set of libraries
  - the ocean Makefile, which compiles the ocean component code together with the IOIPSL and SOFA libraries and produce the final executable binary file, which is stored in the binaries repository
4. change directory to the `../CIGODAS/RUN01/`, where the `sm_SO-OR_TS_CTRL` (see appendix 2) and the `time.txt` are to be set up for the experiment: the former defines all the environment variables and sets the working directories, while the latter contains the current assimilation cycle starting date .
  5. invoke the `qsub` command followed by `sm_SO-OR_TS_CTRL`. The script is able to re-submit itself at the end of each assimilation cycle, until a preset ending date is reached.

The link subroutines between the ocean component and the assimilation component are collected into the `linksub.f` and `interlib.f` files. Such routines provides the transfer of temperature and salinity fields respectively from the ocean code to the assimilation code, and temperature and salinity corrections from the assimilation code to the ocean code. Some procedural entries are inserted into the OPA code via specific include files (`orcamodXX.h` files) activated by the `key_sofa_on` compiling key included into the ocean module Makefile (removing the `key_sofa_on`, a standalone ORCA2 code is recovered). SOFA code calls the OPA code as its own subroutine, so flow control is in

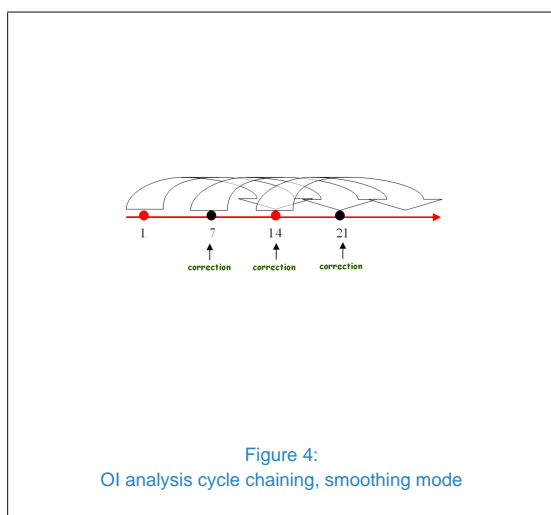
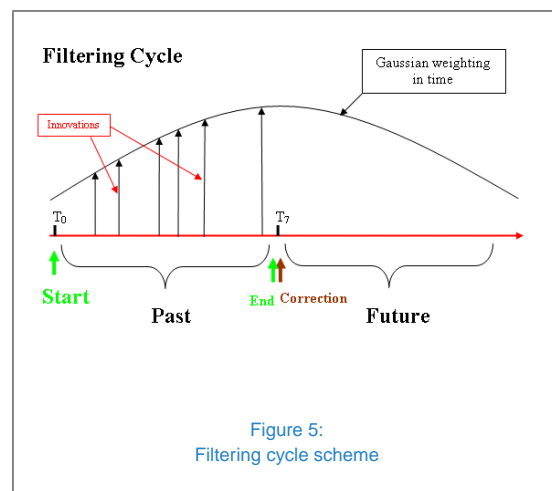
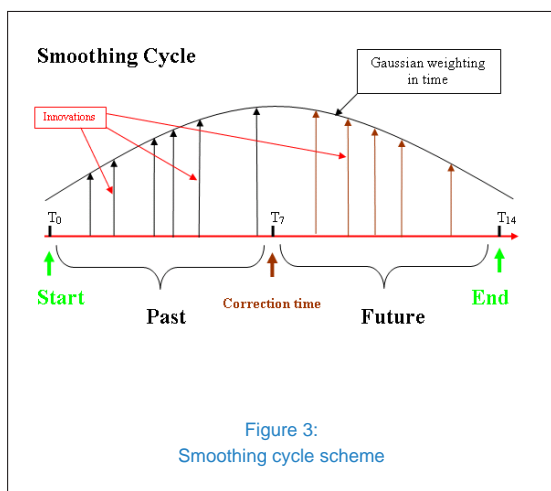
this way demanded to the SOFA. Such an arrangement was chosen to implement two different assimilation methods, called filtering and smoothing. The latter setting needs that the `key_sofa_smooth_emp` is added to the ocean module makefile, otherwise a filtering setup is achieved.

- **Filtering** : with reference to the analysis time, the observed data are gathered only from the past.
- **Smoothing** : with reference to the analysis time, the observed data are gathered both from the past and the future.

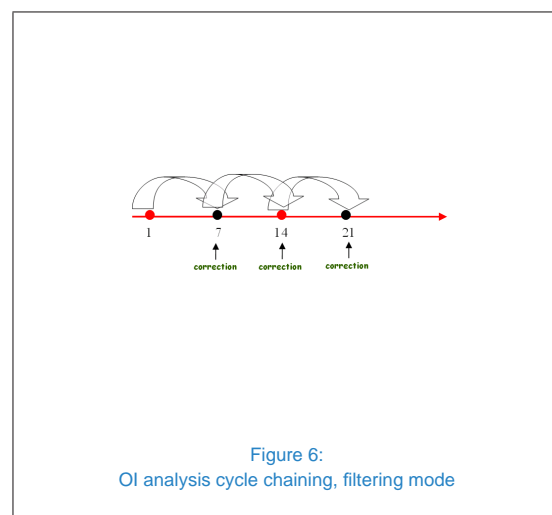
Such methods require a time window for data selection: we choose 7 days for our runs. The typical assimilation cycle is smoothing.

When the smoothing mode is chosen (fig. 3), the OGCM is run for 14 days. During this time interval, innovations (observation-background) are computed using the background state at the appropriate measurement time. A Gaussian weighting function, with a 7-day time scale, is applied to the innovation vector in order to take into account the temporal distance of observations from the time when the model state is actually updated (the analysis time): innovations that are farther (closer) in time from the analysis time will have a lower (larger) weight.

At the end of the OGCM run, the state variables of the model solution at the analysis time (taken as the central time step of the assimilation cycle, when in smoothing), are updated by SOFA with the correction fields computed at each model grid point from the innovation vector. The updated model solution, at the beginning of the 8th day of the assimilation cycle, provides the initial conditions for the following cycle, with the OGCM restarting from the corrected model state (fig. 4).



The filtering mode (fig. 5, fig. 6) is used at the end of the year to recover the 31st December restart date. The innovations are scaled using the same Gaussian weighting function used for smoothing cycles.



The original ROOI (SOFA 3.0) source code has been rearranged: source files, that were originally scattered in a complex directory structure each provided by a local makefile, are now collected together in a common SOURCE directory. A single common Makefile was built accordingly. The source code was also modified to eliminate bugs and improve performances (ROOI1). Subsequently, the code was partially rewritten to implement the gridpoint-defined EOFS (ROOI2).

**ROOI1** : The observation operator, being represented by a constant matrix, is evaluated at the first timestep and stored in memory. This optimization avoids calls to the observation subroutine at every gridpoint, resulting in a better code vectorization and parallelization. The domain splitting eofs (V1) is retained

**ROOI2** : code is based on ROOI1, but more pervasive changes are implemented over the whole code. Several routines have been rewritten to handle eofs defined over each single gridpoint. Basically, the masking technique used to evaluate corrections when using the macro-domain is removed, and direct indexation to the local EOFS is introduced. Computational performance is substantially unchanged from ROOI1

## RUNNING ENVIRONMENT

The experiment repository contains the script *sm\_SO-OR\_TS\_CTRL* (table 1) needed to run a CMCC/INGV-CIGODAS assimilation cycle. At the end of each assimilation cycle, a restart file is produced, which can be used to start the following cycle.

The model is setup to run 14 days of simulation

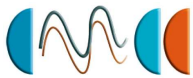
for each assimilation cycle when in smoothing mode. Filtering mode is used to close each year at December 31st.

Information regarding the calendar and cycle length (in days) is put into the *time.txt* file (table 1), which contains year, month, day of the future restart date, the cycle assimilation time window (in days), the year, month and day of the last restart.

Such information are needed both by ocean module and the assimilation module to evaluate the time window, correctly read input files and write the correct dates on output files.

Here we refer to assimilation cycle as a submission of the script covering 14 days of simulation, while the job is intended as a collection of cycles covering the entire experiment time frame. The *sm\_SO-OR\_TS\_CTRL* script, when submitted to the NEC-SX9 by the **qsub** command, performs the following steps:

- Set environment variables related to
  - Input files (table 2, 3, 4, 5) and output files (table 6, 7)
  - Model/experiment name
  - Experiment date and duration
- Modify the namelists according to the experiment setup
  - *namelist.in* for the ocean module (table 10)
  - *sofa.prm* for the assimilation module (table 8, 9)
- Create a scratch subdirectory, where the executables and input files are collected and the model will run
- Launch the model
- Save the output and restart files (table 5, 6) in the storage directories



# 14

- Update the *time.text* file for the next cycle
- Resubmit the *sm\_SO-OR\_TS\_CTRL* to perform the next assimilation cycle, until the end date of the job is reached.

## AUXILIARY FILES

**Table 1**  
**Job script and ancillary file**

Filename	short description
sm.SO-OR.TS.CTRL	script used to launch the experiment
time.txt	ancillary file used to store the starting date of the current assimilation cycle

**Table 2**  
**Ocean component initial conditions**

Filename	short description
EMPave_old.dat	previous cycle E-P correction file <sup>1</sup>
orcaini.nc	previous assimilation cycle restart file <sup>2</sup>

<sup>1</sup> A EMPave\_old.dat file with zero correction is used when starting from scratch

<sup>2</sup> This file is absent when starting from scratch

**Table 3**  
**Ocean component boundary conditions**

Filename	short description
flx_1d.nc	daily heat and freshwater fluxes file
LEVITUS_1m_Salinity_lce_nomask.nc	climatologic monthly salinity file <sup>1</sup>
LEVITUS_1m_Temperature_Pot_lce_nomask.nc	climatologic monthly potential temperature file <sup>1</sup>
runoff_1m_nomask.nc	climatologic river runoff file <sup>2</sup>
taux_1d.nc	daily zonal wind stress file <sup>2</sup>
tauy_1d.nc	daily meridional wind stress file <sup>2</sup>
sst_1d.nc	daily sea surface temperature file <sup>2</sup>

<sup>1</sup> World Ocean Atlas, 1998 <sup>2</sup> ERA40 reanalysis fields

All files in table 2, 3, 4 are , with the exception of the *namelist* file, link files located in the working directory, pointing towards the real input files. This is done to save time, avoiding unnecessary file transfers during execution.

**Table 4**  
**Ocean component miscellaneous files**

Filename	short description
coordinates.nc	model grid coordinate file
dist.coast.nc	distance-to-coast file
namelist	ocean model namelist
ahmcoef	eddy viscosity coefficients file
bathymetry	ocean model bathymetry file

The *ahmcoef* (tab. 4) , *bathymetry* (tab. 4) , *coordinates* (tab. 4) , *dist.coast.nc* (tab. 4) , *LEVITUS* (tab. 3) and *runoff* (tab. 3) files are never changed.

The *flx\_1d.nc* (tab.3) , *sst\_1d.nc* (tab. 3) , *taux\_1d.nc* (tab. 3) and *tauy\_1d.nc* (tab. 3) files are linked to ERA40 daily fluxes, sst and wind stresses. Being each file a collection of daily fields related to a specific year, those files are linked to the appropriate source file every new year.

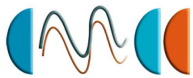
The *EMPave\_old.dat* (tab. 2) and *orcaini.nc* (tab. 2) files are linked to the previous assimilation cycle output files: respectively the *EMPave.dat* (E-P correction, tab.6) and the *restart* () files ( *vp\_EXPERyyyym1d1\_yyyym2d2\_restart.nc* in our example, tab. 6).

**Table 5**  
**Assimilation component input files**

Filename	short description
fort.34	mask of EOF regions file
sofa_corr.adat	correlation radius parameter file
sofa_eofs.dat	EOFs file
sofa.prm	SOFA namelist <sup>1</sup>
sofa_vpS.dat	observed salinity vertical profiles
sofa_vpT.dat	observed temperature vertical profiles

<sup>1</sup> this file is set up inside the run script and generated dynamically

The *fort.34* (tab. 5) , *sofa\_corr.adat* (tab. 5) and *sofa.prm* (tab. 5) files are common in every assimilation cycle.



# 16

The *sofa\_vpS.dat* (tab. 5) and *sofa\_vpT.dat* (tab. 5) files contain the vertical profiles relative to the actual and subsequent month, and are changed monthly. These files are composed by two consecutive months to allow correct misfit calculations in smoothing mode, because the time window usually extends beyond the month end.

The *sofa\_eofs.dat* (tab. 5) file in the SOFA ROOI1 is taken constant, while in the ROOI2 it changes seasonally.



Sample output files for an experiment named EXPER, executed between d1/m1/yyyy and d2/m2/yyyy are given, to illustrate the filename coding (tab. 6)

**Table 6**  
**Ocean component output files<sup>1</sup>**

Filename	short description
date.file	sample output filename file
EMPave.dat	current cycle E-P file
ocean.output	Ocean component log file
solver.stat	elliptic solver log file
time.step	time step counter file
vp.EXPERyyyyy1d1_yyyym2d2_restart.nc	restart file
vp.EXPER_1d_yyyym1d1_yyyym2d2_diagap.nc	diagnostic file
vp.EXPER_1d_yyyym1d1_yyyym2d2_grid.T.nc	T-grid output
vp.EXPER_1d_yyyym1d1_yyyym2d2_grid.U.nc	U-grid output
vp.EXPER_1d_yyyym1d1_yyyym2d2_grid.V.nc	V-grid output
vp.EXPER_1d_yyyym1d1_yyyym2d2_grid.W.nc	W-grid output

<sup>1</sup> Sample output files for an experiment named EXPER, executed between d1/m1/yyyy and d2/m2/yyyy are given, to illustrate the filename coding



## 18

Restart file (tab. 6) contains the corrected restart fields relative to the central time between start and end dates. To allow leapfrog scheme, two consecutive time steps of each needed variable ( temperature, salinity, zonal velocity, meridional velocity, relative vorticity, horizontal divergence) are saved.

The output files contains 2-d and 3-d variables related to the referred grid.

Each output file is stored in NetCDF format. Fields are stored into the output files as a collection of daily means.

Post processing is needed to extract first seven days (or the correct number of days to recover month end) from each file and produce monthly mean fields.

We make use of the ncks ( <http://nco.sourceforge.net/nco.html#ncks-netCDF-Kitchen-Sink> ) and ncra ( <http://nco.sourceforge.net/nco.html> ) utility to extract and mediate fields.

The OI output files (tab. 7) are saved at the end of each assimilation cycle: the *fort.537* is a binary file containing the corrections applied on the Temperature and Salinity fields. The *db.vp* file is a binary file in which the vertical profiles selected for assimilation are stored. The *sofa\_dia\** are diagnostic files containin useful statistics on current cycle assimilation. The *sofa.log* and *infovp\** files are logfiles containing runtime SOFA messages and the general CIGODAS runtime messages, respectively.

Table 7  
Assimilation component output files

Filename	short description
db.vp	vertical profiles file <sup>1</sup>
fort.537	corrections file <sup>1</sup>
infovp_*	assimilation component execution log file
sofa_dia001	temperature assimilation diagnostic file
sofa_dia002	salinity assimilation diagnostic file
sofa.log	assimilation statistics file

<sup>1</sup> binary file

**Table 8**  
**SOFA OI namelists**

Namelist entry	short description
TITLE <sup>1</sup>	The title of the run (string), which will be written in the log file
IDDEBUG <sup>2</sup>	Sets the level of debugging messages written to the log file. Default=0.
NSETS <sup>2</sup>	Number of data sets in input
FNDAT(NSETS) <sup>2</sup>	The names of the files containing data to be analyzed by sofa (string)
SETTYPE(NSETS) <sup>2</sup>	Specifies the data type associated with each dataset. 1 is for vertical profile data (CTD, XBT, etc.)
BOGUS(NSETS) <sup>2</sup>	Specifies the type (if any) of bogus observations associated with each dataset. Default: 0 (normal observations).
VERIF(NSETS) <sup>2</sup>	Says whether datasets contain verification data. 0: data from the dataset will enter the analysis
IUBIAS(NSETS) <sup>2</sup>	Specifies the type of debiasing (if any) applied to data belonging to the specified dataset before diagnostic calculation and analysis. Default=0 no average is removed
IDLOADR(NSETS) <sup>2</sup>	Specifies the data loader which will be used for each dataset with BOGUS=0. The choice of a loader depends on the format and type of data (see SETTYPE), 0 : standard Vertical Profile data loader
AMPMAX <sup>2</sup>	For Type 0 (indirect scalar) data only. The maximum absolute value accepted of the input data
IFREQ <sup>2</sup>	For Type 0 (indirect scalar) data only. The sampling frequency at which to decimate input data points. Default=1.
IHWIN <sup>2</sup>	For Type 0 (indirect scalar) data only. The half-window used to lowpass-filter the input data lines at frequency IFREQ.
IWOLA <sup>2</sup>	if 1, an Off-Line Analysis (OLA) information file will be written.
IDELTA <sup>2</sup>	if 1 : a correction is calculated from model-data misfits
IFRELOAD <sup>2</sup>	Says if we should reload the observational databases of a previous run (logical)
NMASKS <sup>2</sup>	The number of regional masks to be defined for diagnostic calculations.
NRED <sup>3</sup>	The local size of the reduced space, i.e. the number of modes to be estimated by SOFA

<sup>1</sup> id namelist: identification parameters <sup>2</sup> io namelist: input/ output, pre-/post-processing parameters <sup>3</sup> red namelist: order reduction

**Table 9**  
**SOFA OI namelists**

Namelist entry	short description
RGX(NRED), RGY(NRED), RGT(NRED) <sup>4</sup>	The guess error correlation radii for zero time lag in the x- and the y-directions in kilometers and in time.
ROX(NSETS), ROY(NSETS), ROT(NSETS) <sup>4</sup>	The observational error correlation radii for zero time lag in the x- and the y-directions in kilometers and in time.
IFGEOVAR <sup>4</sup>	For geocentric coordinates only. If 0, Covariance radii (RGX, RGY, RGT, ROX, ROY, ROT) and influence radii (RIX, RIY, RIT) do not vary with geographic location. The x and y axes are not rotated.
RIX, RIY, RIT <sup>5</sup>	The selection radii for zero time lag in the x- and the y-directions in kilometers and in time. The selection radii define a perimeter or 'bubble' of elliptic shape in space-time within which the data-model misfits are inverted around each point of the analysis grid.
NSEL(NSETS) <sup>5</sup>	The maximum number of influential points which are to be selected from each dataset for sub-optimal objective analysis.
REDUN(NSETS) <sup>5</sup>	When selecting data in the influence bubble, the suboptimal algorithm eliminates from the selection observations which are closer than $1/\sqrt{REDUN}$ times their distance to the interpolation point.
IFCONV <sup>6</sup>	The type of conversion between the data coordinate system and the analysis/model canvas coordinate system. if 2: data are located in geocentric coordinates (lon,lat).
TTMIN <sup>7</sup>	The starting date of the analysis
NTIMES <sup>7</sup>	The number of integration/analysis cycles.
TINT <sup>7</sup>	The interval between estimates.
IFHCST=0 <sup>7</sup>	Chooses how misfits in the past and future (hindcast) will be handled. 0 is filter mode, 1 is smoother mode.

<sup>4</sup> sta namelist: statistical models <sup>5</sup> sop namelist: suboptimal analysis parameters <sup>6</sup> can namelist: geographical canvas/grid parameters <sup>7</sup> tim namelist: time scheduling parameters

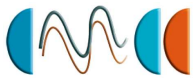


Table 10  
CIGODAS pre-compiling keys

Compiling Key	short description
key_orca_enact	read ENACT input files
key_tradmp	internal restoring to T-S fields
key_orca_r2	ORCA 2 degree mean resolution mesh
key_freesurf_cstvol	free surface arrays
key_diaeiv	diagnostic eddy induced velocities
key_zdftke	tke scheme for vertical mixing
key_fluxsr	penetrative solar radiation option
key_trahdfiso	tracer diffusion along isopycnal surfaces
key_trahdfcoef2d	space varying tracer eddy viscosity coefficient with latitude and longitude
key_trahdfeiv	eddy induced velocity parameterisation (requires key_trahdfiso)
key_dynhdfcoef3d	space varying momentum eddy viscosity coefficient with latitude, longitude and depth
key_convevd	enhanced vertical diffusion
key_sst	prescribed sea surface temperature
key_temdta	read restoring temperature field
key_saldta	read restoring salinity field
key_saldta_monthly	read monthly salinity restoring fields
key_tau	prescribed surface boundary condition on momentum (wind stresses)
key_fix	prescribed heat and freshwater fluxes
key_diagap	compute horizontal averaged and basin averaged diagnostics on tracers
key_ddmixing	double diffusion vertical mixing
key_monotasking	single cpu task
key_diamonthlymean	save exact monthly mean diagnostic fields
key_diamean	save diagnostic mean fields (namelist value in days)
key_diatrdtra	save diagnostics of tracer trends (temperature and salinity)
key_diahth	save depth of the 20°C isotherm
key_smooth_emp	generate correct E-P correction for smoothing mode

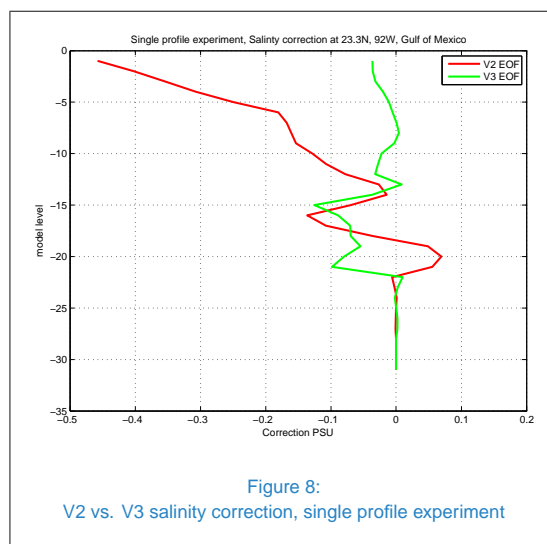
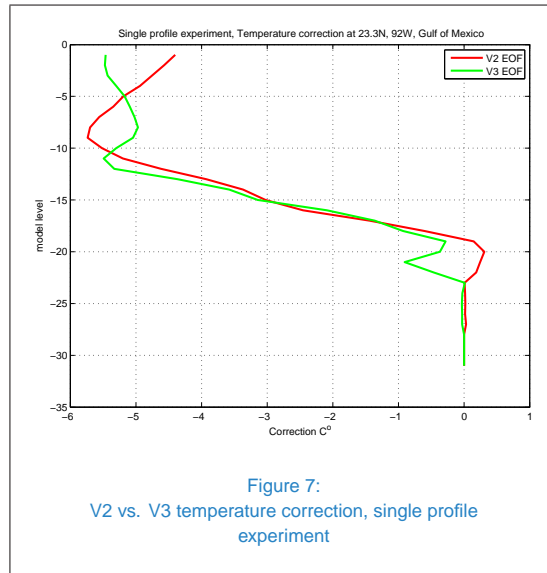
## SINGLE PROFILE ASSIMILATION EXPERIMENT: SENSITIVITY TO EOF

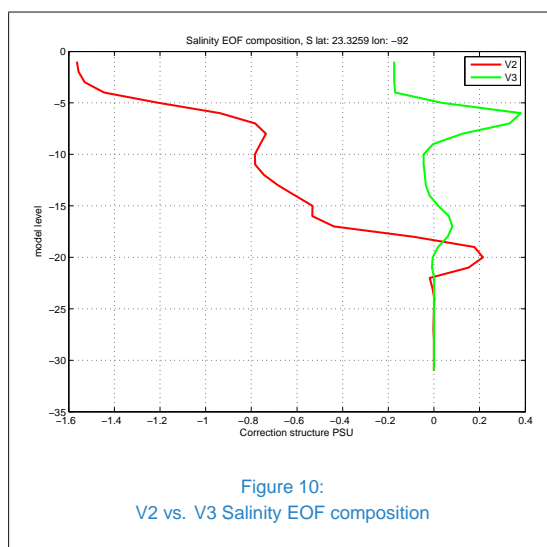
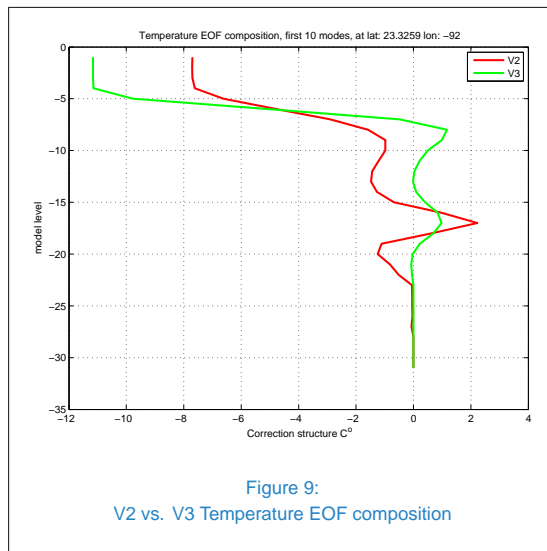
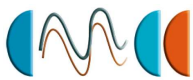
The assimilation results are greatly influenced mainly by the EOFs (vertical statistic). A simple twin single profile (one for temperature and one for salinity) experiment has been performed to test the influence of different EOF sets (V2 and V3). Figure(7) and figure(8) shows the vertical corrections on temperature and salinity produced by the OI, evaluated at 23.3N, 92W inside the Gulf of Mexico. The temperature corrections are roughly similar in magnitude, though differences in shape are evident in the upper layers (level 1 to 10, from 5m to 95m in depth) and around level 20-21 (from 511m to 732m). In the lower layers, the sign of the correction is opposite.

The salinity corrections show similar differences located over the same levels (level 1 to 10 and around level 20-21), but the magnitude of correction in the upper layers is very different.

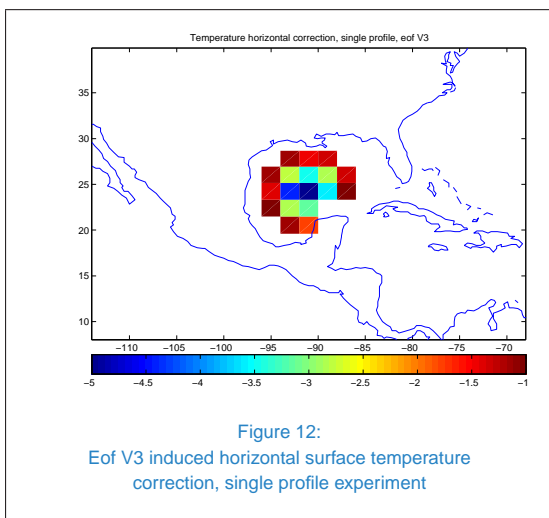
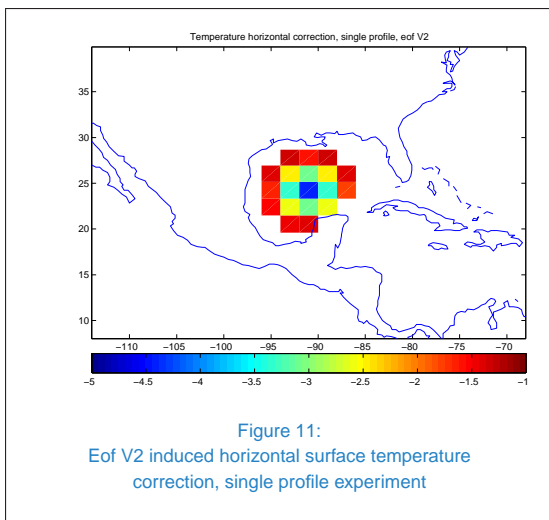
The V2 EOFs set is produced using horizontally smoothed temperature and salinity fields, while the V3 EOFs are produced from un-smoothed fields.

The differences are evident when comparing the EOF composition (first ten modes) at the same location, as showed in figure(9) and (10).



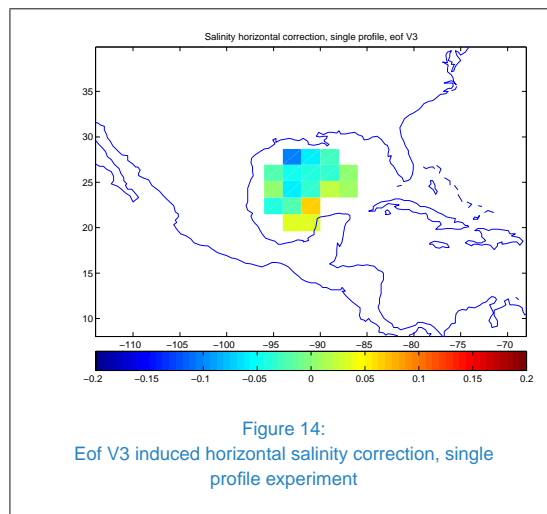
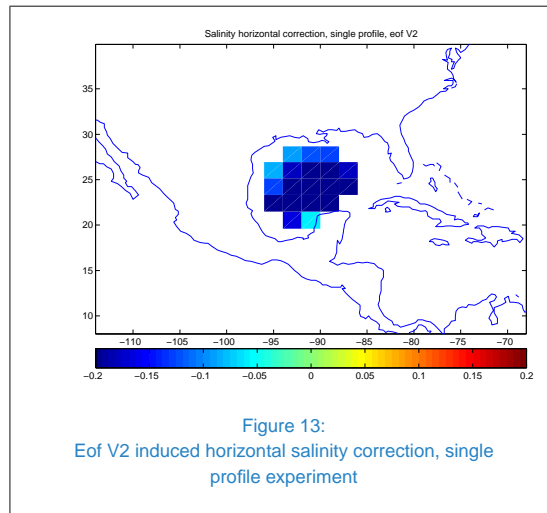


The spatial distribution of corrections induced by the same single profile experiments are showed here. The horizontally gaussian-shaped correction in temperature is evident in both experiments using V2 EOFs (fig. 11) and V3 EOFs (fig.12), where a strong negative correction can be found near the location of the assimilated profile, decreasing in magnitude with distance.



The gaussian shape of salinity corrections at the surface is less evident due to the strong influence of the eofs at each grid point. The local

V2 EOF sets (fig. 13), derived by smoothed T-S fields, produce a more regular pattern in salinity corrections than those of V3 (fig. 14).





## References

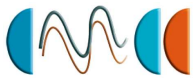
- Bellucci A., Masina S., Di Pietro P., and Navarra A., 2007: Using temperature-salinity relations in a global ocean implementation of a multivariate data assimilation scheme., *Monthly Weather Review*, vol. 135, pp. 3785-3807
- Blanke, B. and Delecluse, P., 1993: Variability of the tropical Atlantic Ocean simulated by a general circulation model with two different mixed-layer physics. *J. Phys. Oceanogr.*, 23, 1363-1388.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- De Mey, P., and M. Benkiran, 2002: A multivariate reduced-order optimal interpolation method and its application to the Mediterranean basin-scale circulation. *Ocean Forecasting: Conceptual Basis and Applications*, N. Pinardi and J. D. Woods, Eds., Springer Verlag, 281-306.
- Ide, K., Courtier, P., Ghil, M., and Lorenc A. C., 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, 75, 181-189
- Ingleby, B., and M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles - historical and real-time data. *Journal of Marine Systems*, 65, 158-175
- Lazar, A., G. Madec, and P. Delecluse, 1999: The deep interior downwelling, the Veronis effect, and mesoscale tracer transport parameterizations in an OGCM. *J. Phys. Oceanogr.*, 29, 2945-2961.
- Levitus, S., and Coauthors, 1998: *World Ocean Database 1998*. NOAA Atlas NESDIS 1, 346 pp.
- Madec G., and M. Imbard, 1996: A global ocean mesh to overcome the North Pole singularity. *Clim. Dyn.*, 12, 381-388.
- Madec, G., Delecluse, P., Imbard, I. and Levy, C., 1999: OPA 8.1 Ocean General Circulation Model reference manual, Note du Pôle de modélisation,, Inst. Pierre-Simon Laplace (IPSL), France, No. 11, 91 pp.
- Sparnocchia, S., Pinardi, N., and Demirov, E., 2003: Multivariate Empirical Orthogonal Function analysis of the upper thermocline structure of the Mediterranean Sea from observations and model simulations, *Ann. Geophysicae*, this issue.
- Troccoli, A., and P. Kallberg, 2004: Precipitation correction in the ERA-40 reanalysis. ERA-40 Project Rep. Series 13, 6 pp.
- UNESCO, 1983: Algorithms for computation of fundamental property of sea water. UNESCO Techn. Paper in Mar. Sci., 44, Unesco, 53pp.



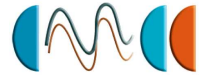
## Appendix 1: OGCM namelist (OPA 8.2)

The OGCM namelist is shown. The parameter values adopted here are used in standard CIGODAS experiments. Pre-compiling keys also are used to select specific parameterizations (see the pre-compiling keys table in the auxiliary files section)

```
:>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
:
: OPA namelist :   model option and parameter input
: -----
:
:
:>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>
:
:
:/::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
:      namrun    parameters of the run
:/::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
:
: no           job number
: cexper       experience name for vairmer format
: lrstar       boolean term for restart (true or false)
: nrstdt        control of the time step (0, 1 or 2)
: nit000       number of the first time step
: nitend       number of the last time step
: ndate0       initial calendar date aammjj
: nbisex       Leap year calendar (0/1)
: nprint       level of print (0 no print)
: nstock       frequency of restart file
: nwrite       frequency of OUTPUT file
: ncpu         number of cpu (for macrotasking run)
: nrunoff      0 no, 1 runoff, 2 runoff+advection
: nwrite =     5475 (annual means)
:
&NAMRUN
```



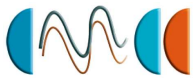
```
no      =      0,
cexper  =  "runname",
lrstar  =  .TRUE.,
nrstdt  =      0,
nit000  =      1,
nitend  =  numts,
ndate0  =  annomesegiorno,
nbisex  =      0,
nprint  =      0,
nstock  =  numts,
nwrite  =  numave,
ncpu    =      1,
nrunoff =      2,
&END
:
:-----
:      namdom    space and time domain (bathymetry, mesh, timestep)
:-----
:
:  ntopo    = 0/1 ,compute/read the bathymetry file
:              (mbathy, nb of T-ocean levels)
:  ngrid    = 0/1, compute/read the horizontal mesh
:              (coordinates, scale factors)
:  nmsh     =1 create a mesh file (coordinates, scale factors, masks)
:  nizoom   index i for the sub domain left bottom
:  njzoom   index j for the sub domain left bottom
:
:  nacc      the acceleration of convergence method
:              = 0, no acceleration, rdt = rdttra
:              = 1, acceleration used, rdt < rdttra(k)
:  atfp     asselin time filter parameter
:  rdt       time step for the dynamics (and tracer if nacc=0)
:  rdtmin    minimum time step on tracers
:  rdtmax    maximum time step on tracers
:  rdth      depth variation of tracer time step
:  nfice     frequency of ice model
:  nclosea  = 0 no closed sea
:              = 1 closed sea (Black Sea, Caspian Sea, Great US Lakes...)
:
&NAMDOM
ntopo    =      1,
ngrid    =      1,
nmsh     =      0,
```



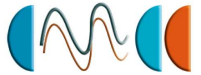
```

nizoom = 1,
njzoom = 1,
nacc = 0,
atfp = 0.1,
rdt = tstp.,
rdtmin = tstp.,
rdtmax = tstp.,
rdth = 800.,
nfice = 6,
nclosea = 1,
&END
:
:-----
:      namhdf    horizontal diffusion
:-----
:
:  ahm0    horizontal eddy viscosity for the dynamics (m2/s)
:  ahmb0    background eddy viscosity for isopycnal diffusion (m2/s)
:  aht0    horizontal eddy diffusivity for tracers (m2/s)
:  ahtb0    background eddy diffusivity for isopycnal diffusion (m2/s)
:  aeiv0    eddy induced velocity coefficient (m2/s)
:
&NAMHDF
  ahm0 = 40000.,
  ahmb0 = 0.,
  aht0 = 2000.,
  ahtb0 = 0.,
  aeiv0 = 2000.,
&END
:
:-----
:      namzdf    vertical diffusion
:-----
:
:  avm0    vertical eddy viscosity for the dynamic (m2/s)
:  avt0    vertical eddy diffusivity for tracers (m2/s)
:  navmt    number of time step for explicit vertical diffusion scheme
:  avevd    vertical coefficient for enhanced diffusion scheme (m2/s)
:  nnpc1    non penetrative convective scheme frequency
:  nnpc2    non penetrative convective scheme print frequency
:  ntrbb1 = 0 no bottom boundary layer on tracers
:           = 1 tracer bottom boundary layer diffusive scheme
:  atrbb1    lateral coefficient for bottom boundary layer scheme(m2/s)

```



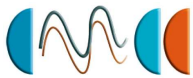
```
&NAMZDF
  avm0   = 1.2e-4,
  avt0   = 1.2e-5,
  navmt  =      3,
  avevd  = 100.,
  nnpc1  =      1,
  nnpc2  =    365,
  ntrbbl =      1,
  atrbbl = 10000.,
&END
:
:-----
:      namric    richardson number dependent vertical diffusion
:                ( #ifdef "key_zdfrichardson" )
:-----
:
:  avmri    maximum value of the vertical viscosity
:  alp      coefficient of the parameterization
:  nric      coefficient of the parameterization
:
&NAMRIC
  avmri = 100.e-4,
  alp   =      5.,
  nric  =      2,
&END
:
:-----
:      namtke    turbulent eddy kinetic dependent vertical diffusion
:                ( #ifdef "key_zdftke" )
:-----
:
:  ediff    coef. to compute vertical eddy coef. (avt=ediff*mxl*sqrt(e) )
:  ediss    coef. of the Kolmogoroff dissipation
:  ebb      coef. of the surface input of tke
:  efave    coef. to applied to the tke diffusion ( avtke=efave*avm )
:  emin     minimum value of tke  $(m^2/s^2)$ 
:  emin0    surface minimum value of tke  $(m^2/s^2)$ 
:  nitke    number of restart iterative loops
:  nmxl     flag on mixing length used
:           = 0 bounded by the distance to surface and bottom
:           = 1 bounded by the local vertical scale factor
:           = 2 first vertical derivative of mixing length bounded by 1
:  npdl     flag on prandtl number
```



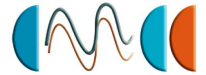
```

:          = 0 no vertical prandtl number (avt=avm)
:          = 1 prandtl number function of richarson number (avt=pdl*avm)
:          = 2 same as = 1 but a shapiro filter is applied on pdl
:
&NAMTKE
  ediff =      0.1,
  ediss =      0.7,
  ebb   =      3.75,
  efave =       1.,
  emin  =     1.e-6,
  emin0 =     1.e-4,
  nitke =       50,
  nmxl  =        2,
  npdl  =        1,
&END
:
:-----
:      namddm    double diffusive mixing parameterization
:-----
:
:  avts      maximum avs for dd mixing
:  hsbfr     heat/salt buoyancy flux ratio
&NAMDDM
  avts  = 1.e-4,
  hsbfr = 1.6,
&END
:
:-----
:      nammbc    lateral and bottom momentum boundary condition
:-----
:
:  shlat     lateral boundary condition on velocity
:            shlat = 0 , free slip
:            0 < shlat < 2 , partial slip
:            shlat = 2 , no slip
:            2 < shlat      , strong slip
:  nbotfr    type of bottom friction
:            nbotfr = 0 , no slip
:            nbotfr = 1 , linear friction
:            nbotfr = 2 , nonlinear friction
:            nbotfr = 3 , free slip
:  bfri1     bottom drag coefficient (linear case)
:  bfri2     bottom drag coefficient (non linear case)

```



```
: bfeb2    bottom turbulent kinetic energy  $(m^2/s^2)$
:
&NAMMBC
  shlat    =      2.,
  nbotfr   =      1,
  bfri1    =    4.e-4,
  bfri2    =    1.e-3,
  bfeb2    =    2.5e-3,
&END
:
:-----
:          namqsr    penetrative solar radiation ('key_flxqsr')
:-----
:
: rabs      fraction of qsr associated with xsil
: xsil      first depth of extinction
: xsil2     second depth of extinction
:
&NAMQSR
  rabs     =    0.58,
  xsil     =    0.35,
  xsil2    =    23.0,
&END
:
:-----
:          namtdp    tracer newtonian damping ('key_tradmp')
:-----
:
: ndmp      type of damping in temperature and salinity
:            (= 'latitude', damping poleward of 'ndmp' degrees and function
:            of the distance-to-coast. Red and Med Seas as ndmp=-1)
:            (= -1 damping only in Med and Red Seas)
: ndmpf     =1 create a damping.coeff NetCDF file (the 3D damping array)
: nmldmp    type of damping in the mixed layer
:            (=0 damping throughout the water column)
:            (=1 no damping in the mixed layer defined by avt >5cm2/s )
:            (=2 no damping in the mixed layer defined rho<rho(surf)+.01 )
: sdmp      surface time scale for internal damping (days)
: bdmp      bottom time scale for internal damping (days)
: hdmp      depth of transition between sdmp and bdmp (meters)
:
: &NAMTDP
: ndmp      =    60,
```



```

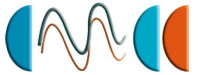
:   ndmpf   =    0,
:   nmltmp  =    1,
:   sdmp    =   50.,
:   bdmp    =  360.,
:   hdmp    =  800.,
: &END
&NAMTDP
:   ndmp    =   -2,
:   ndmpf   =    0,
:   nmltmp  =    0,
:   sdmp    =  1095,
:   bdmp    =  1095,
:   hdmp    =   800,
&END
:
:-----
:   nameos   ocean physical parameters
:-----
:
:   neos     type of equation of state and Brunt-Vaisala frequency
:            = 0, UNESCO (formulation of Jackett and McDougall (1994)
:                  and of McDougall (1987) )
:            = 1, linear: rho(T)   = rau0 * ( 1.028 - ralpha * T )
:            = 2, linear: rho(T,S) = rau0 * ( rbeta * S - ralpha * T )
:   rau0     volumic mass of reference
:   ralpha   thermal expansion coefficient (linear equation of state)
:   rbeta    saline  expansion coefficient (linear equation of state)
:   rcp      ocean specific heat
:
&NAMEOS
:   neos     =    0,
:   rau0     =  1020.,
:   ralpha   =  2.e-4,
:   rbeta    =  0.001,
:   rcp      =  4.e+3,
&END
:
:-----
:   namsol   elliptic solver / island / free surface
:-----
:
:   nsolv    elliptic solver (=1 preconditioned conjugate gradient: pcg)
:            (=2 successive-over-relaxation: sor)

```



```
:                                     (=3 FETI: fet, all require "key_feti" defined)
:  nmax    maximum of iterations for the solver
:  eps     absolute precision of the solver
:  sor     optimal coefficient for sor solver
:  epsisl  absolute precision on stream function solver
:  nmisl   maximum pcg iterations for island
:
:  rnu     strength of the additional force used in free surface b.c.
:
: &NAMSOL
:   nsolv  =      1,
:   nmax   =     800,
:   eps    =   1.E-6,
:   sor    =   1.76,
:   epsisl = 1.e-10,
:   nmisl  =   4000,
:   rnu    =      1.,
: &END
&NAMSOL
  nsolv  =      1,
  nmax   =     800,
  eps    =   1.E-6,
  sor    =   1.92,
  epsisl = 1.e-10,
  nmisl  =   4000,
  rnu    =      1.,
&END
:
: -----
:      namdia    diagnostics on dynamics and/or tracer trends
:                  ('key_diatrdyn' and/or 'key_diatrtra')
:                  or mixed-layer trends ('key_diatrdmld')
: -----
:
:  ntrd    time step frequency dynamics and tracers trends
:  nctls   control surface type in mixed-layer trends (0,1 or n<jpk)
:
&NAMDIA
  ntrd  = 365,
  nfznl = 15,
  ngap  = 15,
  nctls = 0,
&END
```

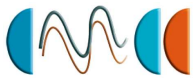




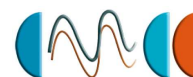
```

:
:-----
:      namspr    surface pressure diagnostic
:-----
:
:  nmaxp    maximum of iterations for the solver
:  epsp     absolute precision of the solver
:  niterp   number of iteration done by the solver
:
&NAMSPR
  nmaxp     =    1000,
  epsp      =    1.e-3,
  niterp    =    400,
&END
:
:-----
:      namcpl    coupled ocean/atmosphere model  (#ifdef "key_coupled")
:-----
:
:  nexco    coupling frequency in time steps
:  cchan    coupling technique 'PIPE' or 'CLIM'
:
&NAMCPL
  nexco          =          24,
  cchan          =      'PIPE',
  nmodcpl        =          2,
  cplmodnam      =      'opa.xx',
  cploasis       =      'Oasis',
  nfldo2c        =          2,
  nflxc2o        =          6,
  ntauc2o        =          4,
  cpl_writ(1)    =      'SOSSTSST',
  cpl_f_writ(1)  =      'ocesst',
  cpl_writ(2)    =      'SOICECOV',
  cpl_f_writ(2)  =      'oceice',
  cpl_readflx(1) =      'SONSFLDO',
  cpl_f_readflx(1) =      'oceflx',
  cpl_readflx(2) =      'SOSHFLDO',
  cpl_f_readflx(2) =      'oceflx',
  cpl_readflx(3) =      'SOTOPRSU',
  cpl_f_readflx(3) =      'oceflx',
  cpl_readflx(4) =      'SOTFSHSU',
  cpl_f_readflx(4) =      'oceflx',

```



```
cpl_readflx(5)   = 'SORUNCOA',
cpl_f_readflx(5) = 'oceflx',
cpl_readflx(6)   = 'SORIVFLU',
cpl_f_readflx(6) = 'oceflx',
cpl_readtau(1)   = 'SOZOTAUX',
cpl_f_readtau(1) = 'ocetau',
cpl_readtau(2)   = 'SOZOTAU2',
cpl_f_readtau(2) = 'ocetau',
cpl_readtau(3)   = 'SOMETAU1',
cpl_f_readtau(3) = 'ocetau',
cpl_readtau(4)   = 'SOMETAU2',
cpl_f_readtau(4) = 'ocetau',
&END
:
:-----
:      namobc      open boundaries parameters (#ifdef key_obc)
:-----
:
:  nieob  i index of the east open boundary
:  niwob  i index of the west open boundary
:  njnob  j index of the north open boundary
:  njsob  j index of the south open boundary
:  rdpeob  time relaxation (days) for the east open boundary
:  rdpwob  time relaxation (days) for the west open boundary
:  rdpnob  time relaxation (days) for the north open boundary
:  rdpsob  time relaxation (days) for the south open boundary
:  zbsic1  barotropic stream function on isolated coastline 1
:  zbsic2  barotropic stream function on isolated coastline 2
:  zbsic3  barotropic stream function on isolated coastline 3
:
&NAMOBC
  nieob = 129,
  niwob = 31,
  njnob = 216,
  njsob = 2,
  rdpein = 1.,
  rdpwin = 1.,
  rdpnin = 30.,
  rdpsin = 1.,
  rdpeob = 1500.,
  rdpwob = 15.,
  rdpnob = 150.,
  rdpsob = 15.,
```



```
zbsic1 = 140e6,  
zbsic2 = 1.e6,  
zbsic3 = 0.,  
&END
```

35



THE CMCC-INGV GLOBAL OCEAN DATA ASSIMILATION SYSTEM (CIGODAS)

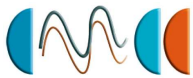
36

Centro Euro-Mediterraneo per i Cambiamenti Climatici

## Appendix 2: Directory tree

Here are listed the CMCC-INGV Global Ocean Data Assimilation System repositories and related most important files:

```
CIGODAS
|-- IOIPSL
|   |-- CVS
|   |-- src
|       |-- obj
|-- SOFA_QU
|   |-- Makefile
|   |-- README-FIRST
|   |-- SOURCE
|       |-- MODIFIED
|   |-- doc
|   |-- profiles
|       |-- arch
|       |-- config
|       |-- config_original
|       |-- default
|       |-- gak
|       |-- io
|       |-- lak
|       |-- misc
|       |-- model
|       |-- sofa
|-- TMP
|-- bin
|-- indata_orca2
|-- lib
|-- src
|   |-- ENSEMBLES_BAK
|   |-- MODIFIED
```



```
| | -- NEW_FILES
| | -- OLDSOURCE
|-- tmp
`-- work
```

■ Ocean component source

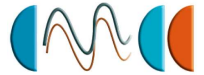
```
.../CIGODAS/src
    common.h
    parameter.h
    parameter.interlib.h
    linksub.f
    Makefile
    Makefile_noSofa
```

■ Ocean input/output module

```
.../CIGODAS/IOIPSL/src
    Makefile
```

■ Assimilation component

```
.../CIGODAS/SOFA_QU/
    Makefile
.../CIGODAS/SOFA_QU/SOURCE/
    Makefile
    orcamod1.h
    orcamod2.h
    orcamod3.h
    orcamod4.h
    orcamod20.h
    orcamod23.h
    interlib.f
    interlib.h
.../CIGODAS/SOFA_QU/profiles/
    config
.../CIGODAS/SOFA_QU/arch/
```



```
.../CIGODAS/SOFA_QU/gak/
.../CIGODAS/SOFA_QU/io/
.../CIGODAS/SOFA_QU/lak/
.../CIGODAS/SOFA_QU/misc/
.../CIGODAS/SOFA_QU/model/
.../CIGODAS/SOFA_QU/sofa/
```

The SOFA namelist is embedded in the running script and generated dynamically

- Compiling scripts

```
.../CIGODAS/work/
    cpl_sofa_qu
    model_setup
```

- Model namelist repository

```
.../CIGODAS/indata_orca2/
    namelist_in
```

- Model executable binaries repository

```
.../CIGODAS/bin/
    orca_opt_codeVERT_PROF_ASSIM_v54_sm
    orca_opt_codeVERT_PROF_ASSIM_v54_fi
```

- Experiment environment

```
.../CIGODAS/RUN01/
    sm_SO-OR_TS_CTRL
    time.txt
```

